May 2000

SECOND-LEVEL IDENTIFICATION (SLID): A TECHNIQUE FOR IMPROVING ID ALGORITHM PERFORMANCE*

Marvin N. Cohen*, Richard Ulrich*, and Adam Hirschel**

*MCA
Atlanta, GA

**Georgia Tech Research Institute
Atlanta, GA

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

ABSTRACT

Second-level identification (SLID), which is a general method that may be applied to reduce errors and improve the performance of any target identification algorithm is presented. In particular, an SLID algorithm using target length for high-range-resolution (HRR) – based identification of aircraft is examined in depth. Results, including improvements over the baseline algorithm performance are given.

INTRODUCTION

The analysis presented in this paper is a part of the work that was done on a Phase 1 SBIR. The purpose of this work was to develop, apply, and evaluate innovative target recognition techniques utilizing a limited but technically challenging train-and-test database of high-range-resolution (HRR) signatures of air targets. The dual goals of the work were to develop and demonstrate techniques that (i) could be of immediate use in improving current-generation identification algorithm performance, while (ii) also laying foundations for the development of next-generation algorithms. These goals were achieved by developing algorithms that were shown to improve the performance of the current generation of techniques and identifying features and new techniques applicable to the next-generation algorithm development effort.

The concept of Second-Level Identification (SLID) algorithms was introduced, developed, and applied. An SLID is a specially designed algorithm that operates on the output of a more general all-target, all-aspect First-Level Identification (FLID) algorithm. The FLID must be designed to handle an "N-class" problem, that is the case where there are many potential target classifications including "unknown" (not in library) and "not declared" (in library, but uncertain as to which). Given typical conditions, any FLID will likely have certain areas of weakness – that is, sets of signatures that are difficult for the FLID to distinguish among – so that target declarations will have a higher than usual chance of being incorrect. The SLID is a tailored special-purpose algorithm designed to address these areas of FLID difficulty, where a subset of the FLID identifications is reassessed using a "2-class" or otherwise-restricted set of possible outcomes.

A conceptual diagram of an SLID structure is given as Figure 1. Focusing on the SLID algorithm that operates on T2 declarations of the FLID, we note that its output can be either T2 or T3. Such an algorithm would be designed and implemented in the case that the FLID tends to identify both T2 and T3 targets as T2 targets. The SLID would then be designed to specifically distinguish T2 from T3 targets, with the desired outcome that the output of the SLID would retain the FLID's efficacy with T2 targets, while correcting many of its misidentifications of T3 targets.

It is shown in this report, for example, that starting with a baseline FLID with good average performance over an entire test dataset, the application of a specially designed SLID at one of its output nodes enabled our system to recover 61% of a particular type of error made by the FLID.

The development of the SLID concept and its application to this particular problem was a part

----

*This work was supported under a Phase I SBIR award to MCA under AFRL, Dayton, Contract # F33615-99-C-1467.

20001204 169

DTIC QUALITY INSPECTED 4

of a broader study undertaken under an SBIR program. Under the program, other features -- such as signature CEPSTRUM; other algorithmic innovations -- such as discriminant pattern processing (DPP); as well as a revolutionary (as opposed to the immediately useful, evolutionary nature of the SLID) HRR target identification approach -- high-accuracy, low-complexity scattering center modeling (HLSCM). (ref 1)

A broader and deeper application of the SLID methodology in conjunction with some of the other techniques developed under the SBIR program should permit the more rapid development of better-performing identification algorithms by allowing the researcher to focus on the features, signal processing, and decision functions that best distinguish certain, particularly difficult classes of targets.

BASELINE DATASET AND (FLID) ALGORITHM

An unclassified six-class HRR signature database and a baseline six-class identification algorithm were acquired and evaluated. This combination of data and algorithm represents a baseline against which improvements may be judged. The database contains 116,697 measured signatures of six distinct target types and a set of synthetically generated signatures at 1 degree pose-angle increments for each target. The baseline algorithm uses single-look forced decision logic based on minimizing the quadratic point-wise distance between each measured signature and the set of synthetic signature templates -- synthetic templates are developed as averages of the synthetic signatures over $5^0 \times 5^0$ target pose sectors. This 'tiling' of the search space is depicted in Figure 3. When applied using these templates and an angle-search of plus or minus one template from the reported angle, the baseline results in Figure 3 are obtained. From the figure it is clear that the algorithm has the most trouble identifying targets of type E. In terms of possible SLID application, the greatest area of opportunity can be seen to exist when a target is declared as a "D".

It is to be noted that the HRR dataset was constructed to be a challenging one that includes target types and signatures that are particularly difficult to identify and separate. On the other hand, the dataset was well-screened to ensure that the data we would be working with would be appropriate for algorithm development and test. It is well-documented (see, for example, ref 2) that corrupted data, both measured and synthetic, are often unknowingly utilized in such studies, making good results difficult to achieve and often leading to misleading conclusions.

The baseline FLID algorithm was designed based on many years of development, experimentation, evaluation, and refinement with these data. While not an 'optimal' algorithm, this algorithm is considered to be quite good for these data. See, for example, ref 3, 4, and 5 which describe multi-look results using a very similar algorithm and train and test datasets.

SLID METHODOLOGY

After baselining the FLID, the measured data were randomly partitioned into developmental and evaluational subsets. The developmental subset was analyzed and utilized repeatedly during the SLID development process. The evaluational subset was never analyzed and was used only as test data to provide final algorithm performance results.

As stated earlier, the primary notion behind the Second Level Identification (SLID) algorithm is that whereas the more general FLID algorithm must be designed to provide the best average performance for, say, a 30-class problem with unknowns, an SLID algorithm can be constructed to perform a much more specific task. By analyzing the specific outputs of the FLID during the algorithm development process and concentrating in those areas in which the FLID evidences its worst performance, we can often reduce the FLID's difficult "N-class" problem to an easier, post-FLID set of SLID "2 or 3 class" problems.

That is, SLIDs can be developed as specialized algorithms designed specifically to excel at solving those problems that are difficult for the broader, more general FLID. One can then view the resulting FLID-SLID system as a sequence of sieves, wherein the FLID is the best general-purpose sieve for the entire problem, while the SLIDs are subsieves designed specifically to correct confusions that are especially difficult for the FLID. An example would be a rule such as the following: "If a target is declared by the FLID as type 'D' with an azimuth between 20 and 50 degrees, then consider carefully whether it might actually be a 'C' or 'E' rather than 'D'."

Indeed, this is exactly the SLID that is reported on herein.

For reasons of scalability, it would be best if the secondary algorithm worked by finding a positive identification trait for one of the alternatives. Less effective but easier and still scalable is if the SLID could identify traits that tend to rule against the initial classification.

Taking an SLID approach is intuitively appealing and experientially justified. That is, it is reasonable to believe that those features and techniques that comprise the FLID – that is, those that provide best overall average performance over a large number of target classes – will not necessarily be the techniques and features best suited to resolve ambiguities between any two or three specific targets from within that set. In the past, to improve an algorithm's performance, we have had to tinker with a single algorithm to try optimize its performance over the entire target set simultaneously. What we have often found is that as we improved performance on, say, Target A, we concomitantly degraded performance on Target B, leaving us with little or no overall average performance improvement. Utilizing an FLID structure, one can retain the algorithm that provides best overall performance, while tailoring specific sub (SLID) algorithms to resolve residual sub-problems.

ESTIMATING TARGET LENGTH

The first specific feature that was evaluated under this contract was related to estimating the electromagnetic length (EML) of a target. EML is expressed as the number of range bins between the beginning and end of the target signature within the HRR signal. Even when the signal-to-noise ratio is very large, there is ambiguity in defining signature regions very precisely. Besides having to establish threshold values and algorithms to specify the beginning and end of each signature, there is the phenomenology of the occasional long "tail" of HRR profiles that (presumably) results from reverberations within cavities of the aircraft.

Two analyses were conducted: one with a simple thresholding technique that attempted to exclude the tails from the EML estimates (*short EML* or, more succinctly, *SEML*) and a second that was specifically designed to include the tails (*long EML* or, more succinctly, *LEML*) in its estimates. It was found that the estimated EML changed with azimuth, per theoretical expectation in both cases. Additionally, it was shown that each of these length estimates (as a function of azimuth) provided utility in distinguishing between the target classes in both measured and synthetic data. Experiments were performed successfully using SEML as an FLID and using LEML as an SLID.

This research suggests that SEML captures much of the performance of the entire HRR profile as an FLID -- a somewhat surprising result, given the time and effort dedicated to designing this HRR-based algorithm. The research further suggests that the LEML feature significantly helps resolve a major residual FLID confusion between targets 'C', 'D', and 'E'.

CLOSEST MATCH METRIC FOR SEML

The first analysis involved a straightforward length estimate for measured signatures based on the mean and standard deviation of the first 100 noise points. Analyses indicated significant – though highly variable – differences in target length between the target classes. The length estimate for the corresponding synthetic template was obtained by applying the same threshold as computed for the measured signature. Using a simple "closest match" metric (i.e., picking the synthetic template whose length is closest to the measured signature), length performed remarkably well as a simple first-level ID mechanism. However, when used as an SLID to resolve the "targets declared as D" problem, this metric degraded rather than improved baseline performance. We concluded that the baseline algorithm already incorporates the information provided by the closest match metric.

LEML ANALYSIS

The second analysis involved looking at the target length using a detection algorithm tuned to find the long signature tails that we believe are due to electromagnetic energy reverberating within target cavities before emerging and reflecting back toward the illuminating radar. In this case, target length proved to be a powerful SLID predictor for the baseline algorithm: measured signatures for target class "D" rarely have target lengths larger than about 90 range bins, whereas C and E (both often misclassified as D) can often evidence longer lengths. An SLID based on this feature was able to recover 61% of the C signatures that had been misclassified as D by the GFE FLID. This

increase in Target C identification came at a "cost" of losing less than 1% of the correctly classified D signatures. A summary of the increased performance gained through the SLID algorithm is given in Figure 4.

Thus, while target extent (length / SEML) performed quite, even surprisingly, well as an FLID, we have found that signature resonances (as estimated by LEML) can also be exploited as an SLID to improve overall performance. Furthermore, these two length measures thus become candidates for a feature-based high-accuracy, low complexity algorithm, which would be the focus of a follow-on Phase II effort.

SUMMARY

Any target identification algorithm will identify some targets with greater accuracy than others. An SLID algorithm allows the weaknesses of the FLID algorithm to be corrected to some extent. Several SLID algorithms may be used on one FLID algorithm to improve accuracy on several different target declarations. Each SLID has the advantage of having to work with only a subset of the general problem, with the opportunity to improve performance on that declaration without affecting the rest of the problem. This is critical because, after significant work has been done on optimizing a given algorithm's performance over an entire dataset, further improving performance on one particular target usually comes at the expense of performance on other targets. Time and effort may be reduced and performance may be significantly improved by utilizing all that the FLID has to offer and then adding SLIDs to improve the weakest of the FLID outputs.

ACKNOWLEDGEMENTS

The authors wish to thank Lt. Anthony Woodcock, AFRL, Dayton, for sponsoring this work under the Phase I SBIR program, Contract # F33615-99-C-1467, and for the SNZ Divisions continuing efforts to sponsor a Phase II follow on. We also express our gratitude to AFRL and Mr. Rob DeWall and other Veridian, Dayton, personnel for providing the baseline GFE data and algorithms. Without this exceptional set of raw material, we could not have made near the progress and contributions we did in such a short time.

REFERENCES

1. "Detection, identification, and correction of errors in measured and synthetic high-resolution databases," Baden and Cohen, Proceedings of the SPIE 11th Annual AeroSense Symposium, Signal Processing, Sensor Fusion, and Target Recognition VI Conference, Advances in Non-Cooperative Target Recognition Session, April 1997.

2. Innovations In Air-To-Air Target Identification, MCA FTR for AFRL, Dayton, Contract # F33615-99-C-1467, Cohen, Ulrich, and Hirschel, May 2000.

3. Post-Processing Options and Their Effects on Target Identification Performance, Proceedings of the IEEE 1998 National Aerospace & Electronics Conference (NAECON), July 1998, Dayton, OH. Cohen and Sylvester.

4. "Performance of a Synthetic-Based Air-To-Air Target Identification Algorithm (U)," Cohen, Baden, Sylvester, and Fox, Proceedings of CISC '97, Secret, San Diego, CA, April 1997.

5. "An Exploration of ATR Discriminant Post-Processing Techniques and How They May Be Combined To Improve Target ID Performance (U)," Sylvester, Cohen, and Fox, Proceedings of the 1997 Tri-Service Radar Symposium, Secret, San Diego, CA, April 1997.

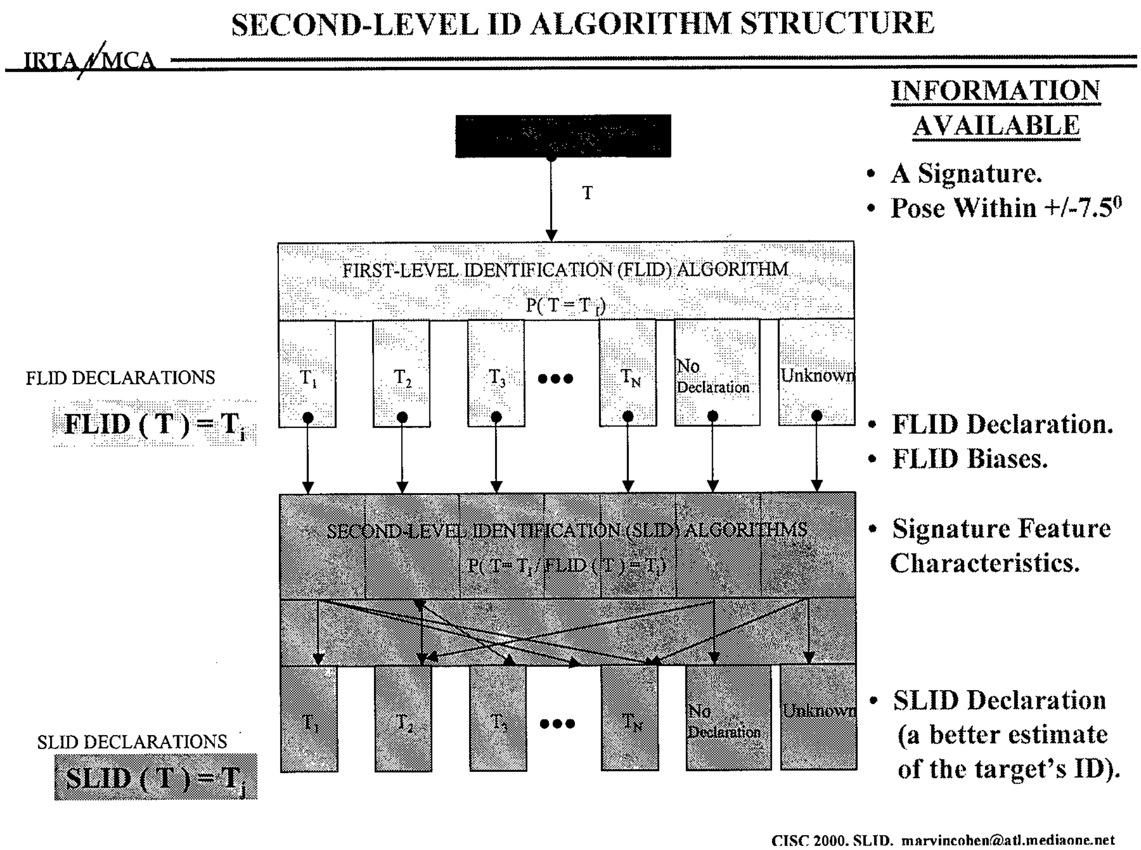

Figure 1 – Second Level ID Concept

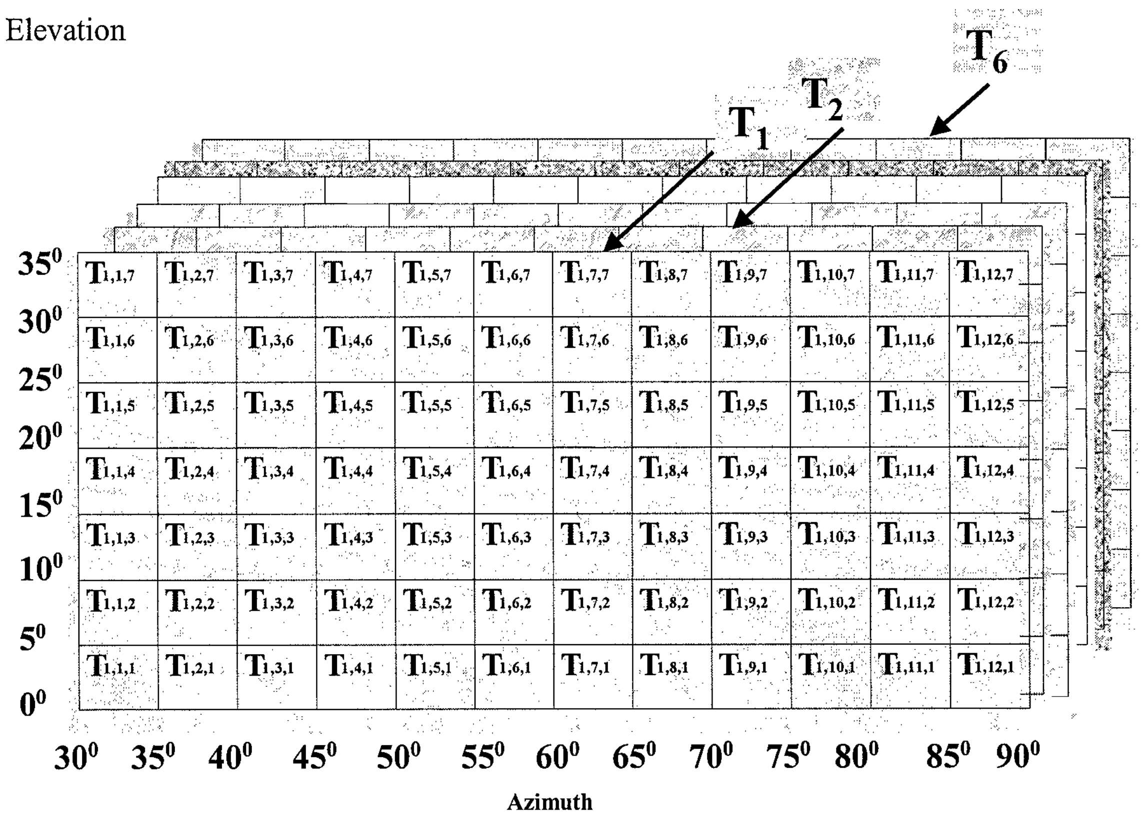

Figure 2 – Tiling the Search Space with Target Templates

BASELINE PROBLEM AREA

IRTA/MCA

	Declared						
Actual	*A*	*B*	*C*	*D*	*E*	*F*	***P(ID)***
A	0.822	0.065	0.047	0.037	0.004	0.025	82%
B	0.040	0.941	0.008	0.004	0.000	0.007	94%
C	0.067	0.025	0.662	0.172	0.017	0.057	66%
D	0.019	0.044	0.032	0.791	0.007	0.107	79%
E	0.049	0.050	0.066	0.141	0.629	0.064	63%
F	0.025	0.029	0.046	0.052	0.042	0.806	81%
P(correct)	80%	82%	77%	66%	90%	75%	

In the baseline algorithm, targets declared as 'D' have a high likelihood of being misclassified.

CISC 2000. SLID. marvincohen@atl.mediaone.net

Figure 3 – Baseline Algorithm Performance

LENGTH-BASED SLID RESULTS

IRTA / MCA

Actual	Declared A	B	C	D	E	F	% recov
A			35	-37	2		0%
B			0	0	0		0%
C			653	-669	16		61%
D			59	-68	9		-0.87%
E			84	-252	168		14%
F			45	-57	12		0%
Total	0	0	876	-1083	207	0	

(was 0.172)

(was 0.141)

Actual	Declared A	B	C	D	E	F	P(ID)
A	0.822	0.065	0.050	0.033	0.004	0.025	82%
B	0.040	0.941	0.008	0.004	0.000	0.007	94%
C	0.067	0.025	0.726	0.106	0.018	0.057	73% (was 66%)
D	0.019	0.044	0.038	0.784	0.008	0.107	78% (was 79%)
E	0.049	0.050	0.075	0.116	0.646	0.064	65% (was 63%)
F	0.025	0.029	0.051	0.046	0.044	0.806	81%
P(correct)	80%	82%	77%	72% (was 66%)	90%	75%	

Average of P(ID) Values 79% (was 78%)

Gross Average P(ID) 79% (was 77%)

CISC 2000. SLID. marvincohen@atl.mediaone.net

Figure 4 – Length SLID Results

DISTRIBUTION STATEMENTS FOR USE ON TECHNICAL DOCUMENTS

A. The following distribution statements and notices are authorized for use on DoD technical documents:

1. DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited.

a. This statement may be used only on unclassified technical documents that have been cleared for public release by competent authority in accordance with DoD Directive 5230.9 (reference (f)). Technical documents resulting from contracted fundamental research efforts will normally be assigned Distribution Statement A, except for those rare and exceptional circumstances where there is a high likelihood of disclosing performance characteristics of military systems, or of manufacturing technologies that are unique and critical to defense, and agreement on this situation has been recorded in the contract or grant.

b. Technical documents with this statement may be made available or sold to the public and foreign nationals, companies, and governments, including adversary governments, and may be exported.

c. This statement may not be used on technical documents that formerly were classified unless such documents are cleared for public release in accordance with reference (f).

d. This statement shall not be used on classified technical documents or documents containing export-controlled technical data as provided in DoD Directive 5230.25 (reference (c)).

2. DISTRIBUTION STATEMENT B. Distribution authorized to U.S. Government agencies only (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office).

a. This statement may be used on unclassified and classified technical documents.

b. Reasons for assigning distribution statement B include:

Foreign Government Information	To protect and limit distribution in accordance with the desires of the foreign government that furnished the technical information. Information of this type normally is classified at the CONFIDENTIAL level or higher in accordance with DoD 5200.1-R (reference (h)).
Proprietary Information	To protect information not owned by the U.S. Government and protected by a contractor's "limited rights" statement, or received with the understanding that it not be routinely transmitted outside the U.S. Government.

Critical Technology	To protect information and technical data that advance current technology or describe new technology in an area of significant or potentially significant military application or that relate to a specific military deficiency of a potential adversary. Information of this type may be classified or unclassified; when unclassified, it is export-controlled and subject to the provisions of DoD Directive 5230.25 (reference (c)).
Test and Evaluation	To protect results of test and evaluation of commercial products or military hardware when such disclosure may cause unfair advantage or disadvantage to the manufacturer of the product.
Contractor Performance Evaluation	To protect information in management reviews, records of contract performance evaluation, or other advisory documents evaluating programs of contractors.
Premature Dissemination	To protect patentable information on systems or processes in the developmental or concept stage from premature dissemination.
Administrative or Operational Use	To protect technical or operational data or information from automatic dissemination under the International Exchange Program or by other means. This protection covers publications required solely for official use or strictly for administrative or operational purposes. This statement may be applied to manuals, pamphlets, technical orders, technical reports, and other publications containing valuable technical or operational data.
Software Documentation	Releasable only in accordance with DoD Instruction 7930.2 (reference (i)).
Specific Authority	To protect information not specifically included in the above reasons and discussions, but which requires protection in accordance with valid documented authority such as Executive Orders, classification guidelines, DoD or DoD Component regulatory documents. When filling in the reason, cite "Specific Authority (identification of valid documented authority)."

3. DISTRIBUTION STATEMENT C. Distribution authorized to U.S. Government Agencies and their contractors (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office).

a. Distribution statement C may be used on unclassified and classified technical documents.

b. Reasons for assigning distribution statement C include:

Foreign Government Information	Same as distribution statement B.
Critical Technology	Same as distribution statement B.
Software Documentation	Same as distribution statement B.
Administrative or Operational Use	Same as distribution statement B.
Specific Authority	Same as distribution statement B.

4. DISTRIBUTION STATEMENT D. Distribution authorized to the Department of Defense and U.S. DoD contractors only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

a. Distribution statement D may be used on unclassified and classified technical documents.

b. Reasons for assigning distribution statement D include:

Foreign Government Information	Same as distribution statement B.
Administrative or Operational Use	Same as distribution statement B.
Software Documentation	Same as distribution statement B.
Critical Technology	Same as distribution statement B.
Specific Authority	Same as distribution statement B.

5. DISTRIBUTION STATEMENT E. Distribution authorized to DoD Components only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

a. Distribution statement E may be used on unclassified and classified technical documents.

b. Reasons for assigning distribution statement E include:

Direct Military Support	The document contains export-controlled technical data of such military significance that release for purposes other than direct support of DoD-approved activities may jeopardize an important technological or operational military advantage of the United States. Designation of such data is made by competent authority in accordance with DoD Directive 5230.25 (reference (c)).
Foreign Government Information	Same as distribution statement B.
Proprietary Information	Same as distribution statement B.
Premature Dissemination	Same as distribution statement D.
Test and Evaluation	Same as distribution statement B.
Software Documentation	Same as distribution statement B.
Contractor Performance Evaluation	Same as distribution statement B.
Critical Technology	Same as distribution statement B.
Administrative/ Operational Use	Same as distribution statement B.
Specific Authority	Same as distribution statement B.

6. DISTRIBUTION STATEMENT F. Further dissemination only as directed by (inserting controlling DoD office) (date of determination) or higher DoD authority.

a. Distribution statement F is normally used only on classified technical documents, but may be used on unclassified technical documents when specific authority exists (e.g., designation as direct military support as in statement E).

b. Distribution statement F is also used when the DoD originator determines that information is subject to special dissemination limitation specified by paragraph 4-505, DoD 5200.1-R (reference (h)).

7. DISTRIBUTION STATEMENT X. Distribution authorized to U.S. Government Agencies and private individuals or enterprises eligible to obtain export-controlled technical data in accordance with reference (c) (date of determination). Controlling DoD office is (insert).

a. Distribution statement X shall be used on unclassified documents when distribution statements B, C, D, E, or F do not apply, but the document does contain technical data as explained in reference (c).

b. This statement shall not be used on classified technical documents; however, it may be assigned to technical documents that formerly were classified.

8. Export Control Warning. All technical documents that are determined to contain export-controlled technical data shall be marked "WARNING - This document contains technical data whose export is restricted by the Arms Export Control Act (Title 22, U.S.C., Sec 2751, *et seq.*) or the Export Administration Act of 1979, as amended, Title 50, U.S.C., App. 2401 *et seq.* Violations of these export laws are subject to severe criminal penalties. Disseminate in accordance with provisions of DoD Directive 5230.25." When it is technically infeasible to use the entire statement, an abbreviated marking may be used, and a copy of the full statement added to the "Notice To Accompany Release of Export Controlled Data" required by DoD Directive 5230.25 (reference (c)).

9. Handling and Destroying Unclassified/Limited Distribution Documents
Unclassified/Limited Distribution documents shall be handled using the same standard as "For Official Use Only (FOUO)" material, and will be destroyed by any method that will prevent disclosure of contents or reconstruction of the document. When local circumstances or experience indicates that this destruction method is not sufficiently protective of unclassified limited information, local authorities may prescribe other methods but must give due consideration to the additional expense balanced against the degree of sensitivity.